

414 Rec'd PCT/PTO 13 NOV 2000

09/700311 ✓

FORM PTO-1390 REV. 5-93		US DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEYS DOCKET NUMBER P00,1796
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371			U.S. APPLICATION NO. (if known, see 37 CFR 1.5)
INTERNATIONAL APPLICATION NO. PCT/DE99/01308	INTERNATIONAL FILING DATE 03 MAY 1999	PRIORITY DATE CLAIMED 11 MAY 1998	
TITLE OF INVENTION METHOD AND ARRANGEMENT FOR DETERMINING SPECTRAL SPEECH CHARACTERISTICS IN A SPOKEN EXPRESSION			
APPLICANT(S) FOR DO/EO/US MARTIN HOLZAPFEL			
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:			
1	<input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.		
2	<input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.		
3	<input checked="" type="checkbox"/> This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay.		
4	<input checked="" type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.		
5	<input checked="" type="checkbox"/> A copy of International Application as filed (35 U.S.C. 371(c)(2)) - drawings attached.		
6	<input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau). <input type="checkbox"/> has been transmitted by the International Bureau. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US)		
7	<input checked="" type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)) - drawings attached.		
8	<input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. §371(c)(3))		
9	<input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau). <input type="checkbox"/> have been transmitted by the International Bureau. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. <input checked="" type="checkbox"/> have not been made and will not be made.		
10	<input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).		
11	<input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).		
12	<input checked="" type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).		
Items 11. to 16. below concern other document(s) or information included:			
11	<input checked="" type="checkbox"/> An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98; (PTO 1449, Prior Art, Search Report).		
12	<input checked="" type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 C.F.R. 3.28 and 3.31 is included. (SEE ATTACHED ENVELOPE)		
13	<input checked="" type="checkbox"/> Amendment "A" Prior to Action. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.		
14	<input type="checkbox"/> A substitute specification.		
15	<input checked="" type="checkbox"/> A change of address letter attached to the Declaration.		
16	Other items or information:		
	<input checked="" type="checkbox"/> Request for Approval of Drawing Additions, 3 sheets of drawings, Figures 1-5. <input checked="" type="checkbox"/> Appointment of Associate Power of Attorney <input checked="" type="checkbox"/> EXPRESS MAIL #EL655299400US dated November 13, 2000		

U.S. APPLICATION NO. (if known) (see 37 C.F.R. 1.51)

09/700311

INTERNATIONAL APPLICATION NO.

PCT/DE99/01308

ATTORNEY'S DOCKET NUMBER

P00,1796

09/700311

17. ☒ The following fees are submitted:**BASIC NATIONAL FEE (37 C.F.R. 1.492(a)(1)-(5):**

Search Report has been prepared by the EPO or JPO \$860.00

International preliminary examination fee paid to USPTO (37 C.F.R. 1.482) .. \$690.00

No international preliminary examination fee paid to USPTO (37 C.F.R. 1.482) but
international search fee paid to USPTO (37 C.F.R. 1.445(a)(2)) \$710.00Neither international preliminary examination fee (37 C.F.R. 1.482) nor international
search fee (37 C.F.R. 1.445(a)(2)) paid to USPTO \$1000.00International preliminary examination fee paid to USPTO (37 C.F.R. 1.482) and all
claims satisfied provisions of PCT Article 33(2)-(4) \$ 100.00**ENTER APPROPRIATE BASIC FEE AMOUNT =**

CALCULATIONS

PTO USE ONLY

\$ 860.00

Surcharge of \$130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months
from the earliest claimed priority date (37 C.F.R. 1.492(e)).

\$

Claims	Number Filed	Number Extra	Rate
Total Claims	10 - 20 =	0	X \$ 18.00
Independent Claims	03 - 3 =	0	X \$ 80.00
Multiple Dependent Claims			\$270.00 +
TOTAL OF ABOVE CALCULATIONS =			\$ 860.00
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 C.F.R. 1.9, 1.27, 1.28)			\$
SUBTOTAL =			\$ 860.00
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).			\$
TOTAL NATIONAL FEE =			\$ 860.00
Fee for recording the enclosed assignment (37 C.F.R. 1.21(h). The assignment must be accompanied by an appropriate cover sheet (37 C.F.R. 3.28, 3.31). \$40.00 per property			+
TOTAL FEES ENCLOSED =			\$ 860.00
			Amount to be refunded \$
			charged \$

a. ☒ A check in the amount of \$ 860.00 to cover the above fees is enclosed.b. ☐ Please charge my Deposit Account No. _____ in the amount of \$ _____ to cover the above fees.
A duplicate copy of this sheet is enclosed.c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any
overpayment to Deposit Account No. 501519. A duplicate copy of this sheet is enclosed.NOTE: Where an appropriate time limit under 37 C.F.R. 1.494 or 1.495 has not been met, a petition to revive (37 C.F.R. 1.137(a) or (b)) must be
filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

SCHIFF HARDIN & WAITE
PATENT DEPARTMENT
6600 Sears Tower
233 South Wacker Drive
Chicago, Illinois 60606-6473

SIGNATURE

Mark Bergner

NAME

45,877

Registration Number

529 Rec'd PCT/PTC 13 NOV 2000

BOX PCT
IN THE UNITED STATES DESIGNATED/ELECTED OFFICE
OF THE UNITED STATES PATENT AND TRADEMARK OFFICE
UNDER THE PATENT COOPERATION TREATY--CHAPTER II

10 Assistant Commissioner for Patents,
Washington, D.C. 20231

Sir:

IN THE SPECIFICATION:

On substitute page 1:

replace lines 1-2, with

20 --SPECIFICATION

TITLE

METHOD AND ARRANGEMENT FOR DETERMINING SPECTRAL SPEECH CHARACTERISTICS IN A SPOKEN EXPRESSION

BACKGROUND OF THE INVENTION

25 Field of the Invention—;

above line 5, insert

--Description of the Related Art--;

in line 6, cancel "thereby";

in line 7, , after "ear", insert –from these sounds–;

in line 8, replace ". In particular, the sounds are thereby" with –, particularly where these sounds are–;

in line 10, replace "[1]" with –I. Daubechies, "Ten Lectures on Wavelets", Saim Verlag, 1992, ISBN 0-89871-274-2, Ch. 5.1, pp. 129-137–;

5 in line 12, replace ". A" with –, resulting in a–;

in line 13, cancel ", respectively,", and cancel "thereby";

in line 14, cancel "ensues", and cancel "English art term:";

in line 16, replace "US-A-5528725" with –U.S. Patent No. 5,528, 725–;

in line 18, before "EP", insert –European Patent–;

10 above line 21, insert

--SUMMARY OF THE INVENTION--;

replace lines 24-25 with

-- This object is achieved by a method for determining spectral speech characteristics in a spoken expression, comprising the steps of: a) digitizing the expression; b) wavelet transforming the digitized expression; and c) defining speaker-specific characteristics based on different transformation stages of the wavelet transformation. –;

15

in line 26, replace "A method" with –The invention provides a method–;

and

20 in line 27, cancel "is recited in the scope of the invention".

On page 2:

in line 1, cancel "thereby";

in line 2, after the last "filter", insert –,–;

in line 3, cancel ", respectively,";

25 in line 5, cancel ", respectively,";

in line 7, replace "whereby" with –where–;

in line 8, after "i.e.", insert –,–;

in line 13, cancel "thereby";

in line 14, cancel "comprised therein";

in line 15, after "i.e.", insert --;
in line 18, cancel "comprised therein";
in line 22, cancel "comprised therein";
in line 23, replace "be defined in" with --is defined such--; and
in line 25, replace "passes" with --pass parts--.

5

On page 3:

in line 3, after the first "as", insert --a--;
in line 5, after "without", insert --a--;
in line 6, after "as", insert --a--;
in line 13, replace "as" with --is--;
in line 23, replace "a0" with --a)--; and
in line 27, cancel "thereby--.

10

On page 4:

in line 8, after "example", insert --,--;
in line 9, cancel "comprised therein";
in line 11, replace "Furthermore" with --Further--;
in line 12, replace "representation" with --representations--; and before the
last "the", insert --so--;

15

in line 14, before "can", insert --so they--;
in line 16, replace "-- loss-free" with --without loss--;
in line 17, replace "Further" with --Furthermore--;
in line 18, replace "recited" with --provided--;
in line 19, replace "digitalized" with --digitized--;
replace line 24 with

20

-- Advantageous embodiments include adding a step to the inventive
method of implementing a windowed transformation of the digitalized expression
into a frequency domain before the wavelet transformation, which may be
implemented with a fast Fourier transformation. An advantageous embodiment

25

may also include a step of determining a low-pass part and a high-pass part of a signal to be transformed in each stage of the wavelet transformation. The high pass part can be subdivided into a real part and an imaginary part.

In the inventive method, the wavelet transformation may include a plurality of transformation stages, a last transformation stage of the plurality of transformation stages supplying a constant part of the expression in a repeated low-pass filtering corresponding to the plurality of transformation stages. Speaker-specific characteristics may be determined by: a basic frequency of the spoken expression; spectral envelope; and/or a huskiness of the spoken expression, and individual speaker-specific characteristics may be adapted to provide a natural sounding concatenation of speech sounds.

An inventive method may be provided implementing the above method for determining spectral speech characteristics comprising a step of selecting those speech sounds from a predetermined data set that assure a natural sounding concatenation of speech sounds on a basis of individual the spectral speech characteristics.

Finally, the object of the invention may be achieved with an arrangement for determining spectral speech characteristics in a spoken expression, comprising a processor unit that is configured to digitize the expression, wavelet transform the digitized expression, and define speaker-specific characteristics on a basis of different transformation stages of the wavelet transformation. --

above line 25, insert

--BRIEF DESCRIPTION OF THE DRAWINGS --;

cancel line 27;

in line 28, before "a wavelet", insert --is a graph illustrating--; and
in line 29, before "a wavelet", insert --is a graph illustrating--.

On page 5:

in line 1, before "a cascaded", insert --is a block diagram illustrating--;

in line 3, replace "Figure 4" with --Figures 4A-4F are graphs illustrating

frequency spectrums of-;

in line 4, before "steps", insert -are pictorial diagrams illustrating the-
above line 5, insert

--DESCRIPTION OF THE PREFERRED EMBODIMENTS--;

5 replace line 6 with -where-;

in line 13, before "imaginary", insert -the-;

in line 17, replace "whereby" with -where-;

in line 19, after "high-pass", insert -part/filter-, after "low-pass", insert
-part/filter-, and replace ".", with -, producing-; and

10 in line 20, replace "In" with -in-, and replace "thereby occurs, i.e." with
-, i.e.,-.

On page 6:

in line 2, replace "304" with -302-;

in line 3, after "pass", insert -part-;

15 in line 4, replace "Im 1" with -Im1-; and

in line 6, after all instances of "pass", insert -part-.

On page 7:

in line 2, replace "Mi" with -Min-;

in line 3, replace "Said" with -These-;

20 in lines 4-5, replace "thereby of particular significance" with
-particularly significant-;

in line 8, cancel "thereby";

in line 9, cancel "comprised";

in line 10, replace "whereby" with -where-;

25 in line 12, after "With", insert -a-;

in line 16, after "shows" insert -the-;

in lines 17-18, replace ", whereby" with -in which-;

in line 19, replace "representatives" with -representations-;

in line 22, cancel "all the"; and
in line 25, replace "thereby to be" with -also-, and replace "dependent"
with -depending-.

On page 8:

- 5 in line 2, replace "ensue be adaptation" with -be implemented by
adapting-;
- in line 4, replace ", whereby" with -in which-, and replace "represent"
with -are-;
- in line 8, cancel "respectively", and replace "whereby" with -where-;
- 10 in line 9, replace "whereby" with -where-;
- in line 16, replace "is comprised in" with -has-;
- in line 17, replace ". However, said" with -which are-;
- in line 18, replace "thereby occur. When, in" with -. In-, and replace
"[...]" with -so-;
- 15 in line 19, replace "whereby" with -where-;
- in line 21, before "gradual", insert -a-;
- replace lines 25-27 with
- The above-described apparatus and method are illustrative of the
principles of the present invention. Numerous modifications and adaptations
20 thereof will be readily apparent to those skilled in this art without departing from
the spirit and scope of the present invention.--.

IN THE CLAIMS:

On page 9:

- replace line 1 with --WHAT IS CLAIMED IS:--;
- 25 Please amend claims 1-10 as follows:
1. (Amended) A method [Method] for determining spectral speech
characteristics in a spoken expression, comprising the steps of:
- a) digitizing said [whereby the] expression [is digitalized];

- b) wavelet transforming said digitized [whereby the digitalized] expression [is subjected to a wavelet transformation]; and
- c) defining [whereby the] speaker-specific characteristics based on [are defined on the basis of] different transformation stages of said [the] wavelet transformation.

5

2. (Amended) The method [Method] according to claim 1, further comprising the step of implementing [whereby] a windowed transformation of said [the] digitalized expression into a frequency domain [is implemented] before said [the] wavelet transformation.

10

3. (Amended) The method [Method] according to claim 2, wherein said step of implementing said windowed transformation is implemented [whereby the transformation into the frequency domain is implemented] with a fast Fourier transformation.

15

4. (Amended) The method [Method] according to claim 1, further comprising the step of: [one of the preceding claims, whereby] determining a low-pass part and a high-pass part of a signal to be transformed [are determined] in each stage of said [the] wavelet transformation.

20

5. (Amended) The method [Method] according to claim 1, further comprising the step of: [one of the preceding claims, whereby] subdividing a high-pass part into [is subdivided according to] a real part and an imaginary part.

6. (Amended) The method [Method] according to claim 1, wherein said step of wavelet transformation further comprises [one of the preceding claims, whereby the wavelet transformation comprises] a plurality of transformation

stages, a [whereby the] last transformation stage of said plurality of transformation stages supplying [supplies] a constant part of said [the] expression in a repeated low-pass filtering corresponding to said [the] plurality of transformation stages.

- 5 7. (Amended) The method [Method] according to claim 1 [one of the preceding claims], wherein said [whereby the] speaker-specific characteristics are determined by an attribute selected from the group consisting of
- a) a basic frequency of the spoken expression;
- b) spectral envelope; and
- 10 c) a huskiness of the spoken expression.

 8. (Amended) The [Employment of the] method according to claim 1, further comprising the step of [one of the claims 1 through 7 for speech synthesis, whereby] adapting individual speaker-specific characteristics [are adapted in view of] to provide a natural sounding concatenation of speech sounds.

- 15 9. (Amended) A [Employment of the] method for implementing the method according to claim 1, comprising the step of: [according to one of the claims 1 through 7 for speech synthesis, whereby]
- selecting those speech sounds from a predetermined data set that assure a natural sounding concatenation of speech sounds [are selected] on a [the] basis of
- 20 individual said spectral speech characteristics.

 10. (Amended) An arrangement [Arrangement] for determining spectral speech characteristics in a spoken expression, comprising:

- a processor unit that is configured to digitize said expression, wavelet transform said digitized expression, and define speaker-specific characteristics on
- 25 a [such that the following steps can be implemented:

- a) the expression is digitalized;
- b) the digitalized expression is subjected to a wavelet transformation;
- c) the speaker-specific characteristics are defined on the] basis of different transformation stages of the wavelet transformation.

5 **IN THE ABSTRACT:**

On page 11:

cancel lines 2-3; and


in line 5, replace "whereby" with --where--.

REMARKS

10 The present Amendment revises the specification and claims to conform to United States patent practice, before examination of the present PCT application in the United States National Examination Phase. All of the changes are editorial and applicant believes no new matter is added thereby. The amendment of claims 1-10 is not intended to be a surrender of any of the subject
15 matter of those claims.

Early examination on the merits is respectfully requested.

Submitted by,

 (Reg. No. 45,877)

20

Mark Bergner
Schiff Hardin & Waite
Patent Department
6600 Sears Tower
233 South Wacker Drive
Chicago, Illinois 60606-6473
25 (312) 258-5779
Attorneys for Applicant

-1-

BOX PCT
IN THE UNITED STATES DESIGNATED/ELECTED OFFICE
OF THE UNITED STATES PATENT AND TRADEMARK OFFICE
UNDER THE PATENT COOPERATION TREATY--CHAPTER II

5 APPLICANT(S): MARTIN HOLZAPFEL
 ATTORNEY DOCKET NO.: P00,1796
 INTERNATIONAL APPLICATION NO: PCT/DE99/01308
 INTERNATIONAL FILING DATE: 03 MAY 1999
 INVENTION: METHOD AND ARRANGEMENT FOR DETERMINING
 SPECTRAL SPEECH CHARACTERISTICS IN A SPOKEN
 EXPRESSION

10 Assistant Commissioner for Patents,
 Washington, D.C. 20231

REQUEST FOR APPROVAL OF DRAWING ADDITIONS

Sir:

15 Enclosed are 3 sheets of drawings, Figures 3-5, showing in red, the
 addition of labels to the elements depicted therein. Approval of the additions is
 respectfully requested.

Submitted by,

20 Mark Bergner (Reg. No. 45,877)
 Mark Bergner
 SCHIFF HARDIN & WAITE
 PATENT DEPARTMENT
 6600 Sears Tower
 Chicago, Illinois 60606-6473
 (312) 258-5779
25 Attorney for Applicant(s)

1/3
FIG 1

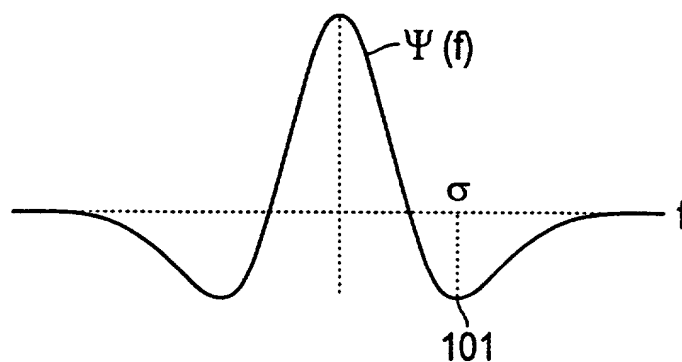


FIG 2

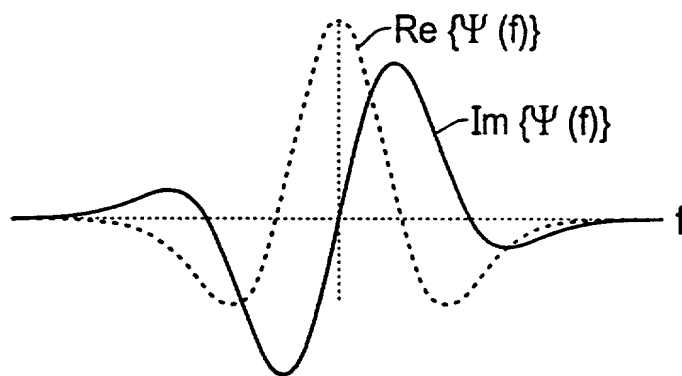
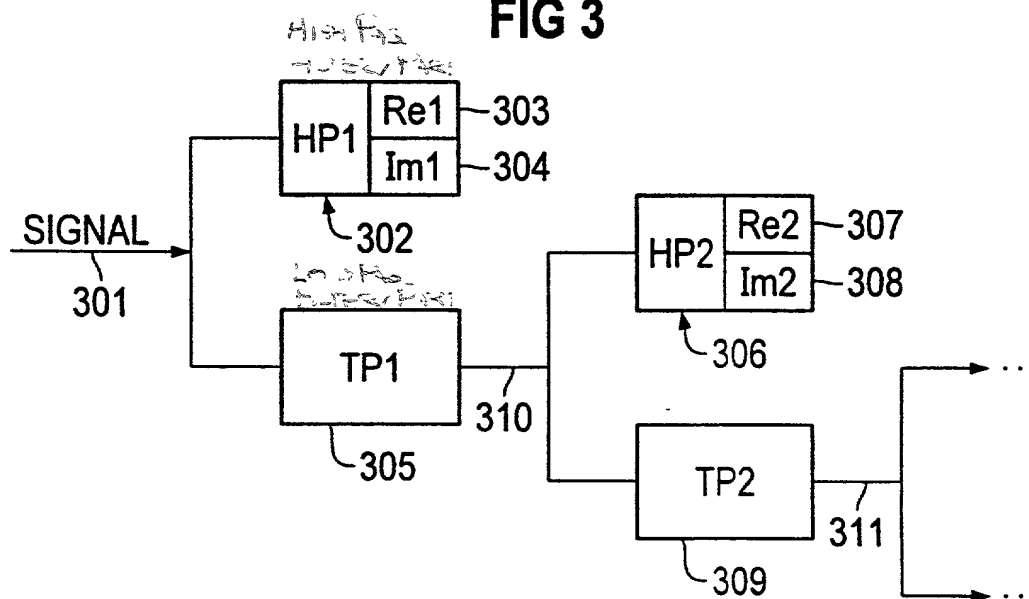
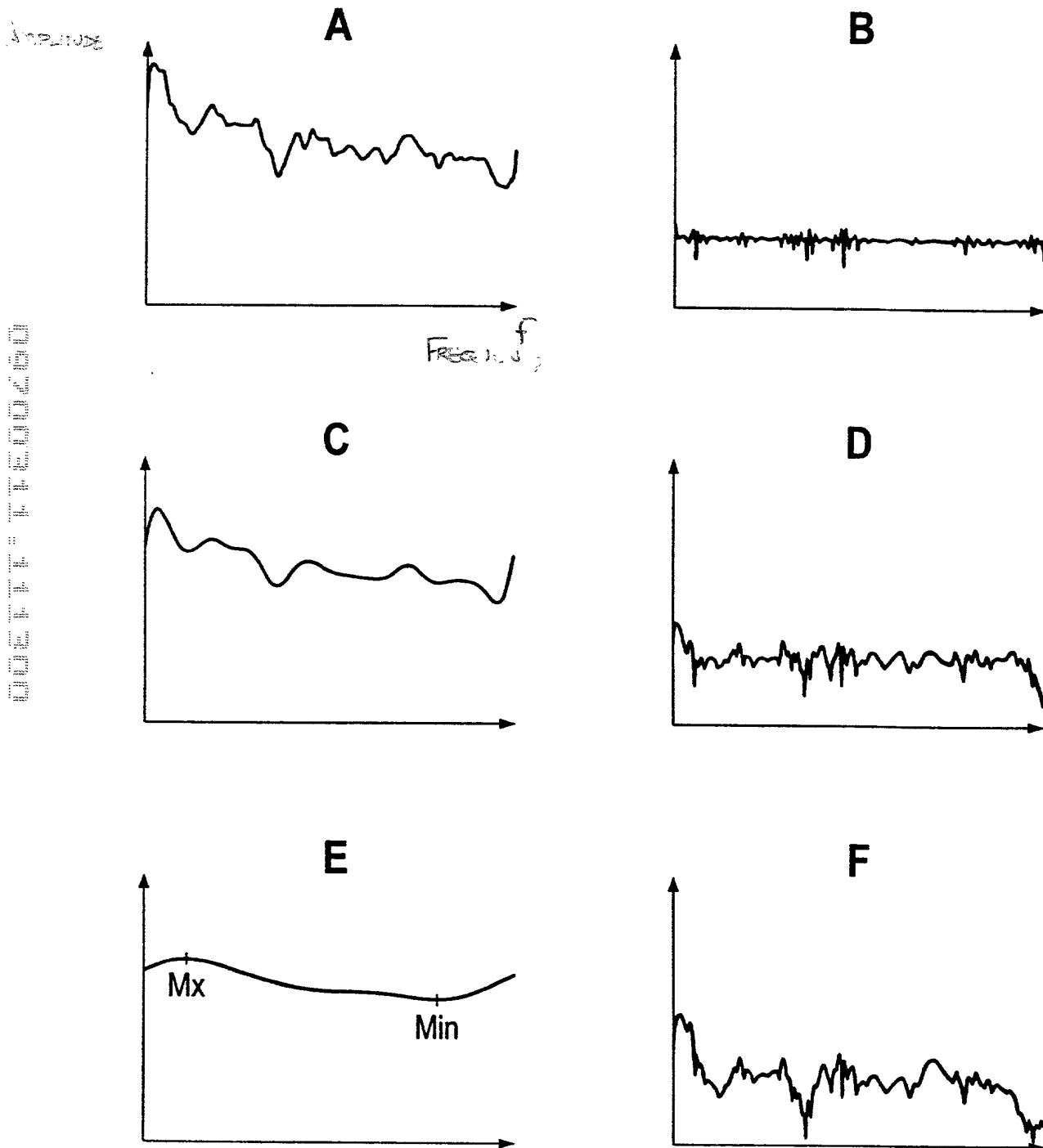


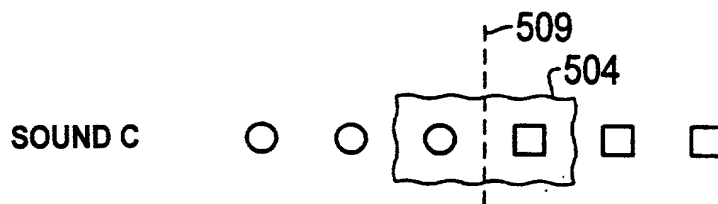
FIG 3



2/3

FIG 4





**METHOD AND ARRANGEMENT FOR DETERMINING SPECTRAL
SPEECH CHARACTERISTICS IN A SPOKEN EXPRESSION**

The invention is directed to a method and to an arrangement for determining spectral speech characteristics in a spoken expression.

5 In a concatenative speech synthesis, individual sounds are combined from speech data banks. In order to thereby obtain a speech curve that sounds natural to the human ear, discontinuities must be avoided at the points where the sounds are combined (concatenation points). In particular, the sounds are thereby phonemes of a language or a combination of a plurality of phonemes.

10 [1] discloses a wavelet transformation. In wavelet transformation, a wavelet filter assures that a respective high-pass part and low-pass part of a following transformation stage completely restore a signal of a current transformation stage. A reduction of the resolution of the high-pass part or, respectively, low-pass part thereby ensues from one transformation stage to the next (English art term: "sub-sampling").
15 In particular, the plurality of transformation stages is finite due to the sub-sampling.

US-A-5528725 discloses a method for speech recognition with wavelet transformations.

EP-A-0519802 discloses a method for speech synthesis that adapts speaker-specific characteristics in view of a natural sounding concatenation of speech
20 sounds.

The object of the invention is comprised in specifying a method and an arrangement for determining spectral speech characteristics with whose assistance, in particular, a speech output that sounds natural can be determined.

This object is achieved according to the features of the independent
25 claims.

A method for determining spectral speech characteristics in a spoken expression is recited in the scope of the invention. To that end, the spoken expression is digitalized and subjected to a wavelet transformation. The speaker-specific characteristics are determined on the basis of different transformation stages of the
30 wavelet transformation.

One advantage, in particular, is thereby that the expression is divided in the wavelet transformation with a high-pass filter and a low-pass filter and different high-pass parts or, respectively, low-pass parts of different transformation stages contain speaker-specific characteristics.

5 The individual high-pass parts or, respectively, low-pass parts of different transformation stages stand for predetermined speaker-specific characteristics, whereby both high-pass part as well as low-pass part of a respective transformation stage, i.e. the respective characteristic, can be modified separately from other characteristics. When, in inverse wavelet transformation, the original signal is in turn
10 combined from the respective high-pass and low-pass parts of the individual transformation stages, then it is assured that it is exactly the desired characteristic that has been modified. It is thus possible to modify certain predetermined peculiarities of the expression without the rest of the expression being thereby influenced.

One development is comprised therein that the expression is windowed
15 before the wavelet transformation, i.e. a predetermined set of samples are cut out, and is transformed into the frequency domain. In particular, a fast-Fourier-Transformation (FFT) is employed for this purpose.

A further development is comprised therein that a high-pass part of a transformation stage is split into a real part and an imaginary part. The high-pass part
20 of the wavelet transformation corresponds to the difference signal between the current low-pass part and the low-pass part of the preceding transformation stage.

In particular, one development is comprised therein that the number of transformation stages of the wavelet transformation to be implemented be defined in that a constant part of the expression is contained in the last transformation stage,
25 which is composed of series-connected low-passes. The signal as a whole can then be presented by its wavelet coefficients. This corresponds to the complete transformation of the information of the signal excerpt into the wavelet space.

When, in particular, only the respective low-pass part is further-transformed (with a high-pass and a low-pass filter), then the difference signal remains as high-pass part of a transformation stage, as explained above. When difference signals (high-pass parts) are accumulated over the transformation stages, then the information of the spoken expression without constant part is obtained in the last transformation stage as cumulative high-pass part.

In the scope of an additional development, the speaker-specific characteristics can be identified as:

a) Basic Frequency:

The oscillation of the high-pass part of the first or of the second transformation stage of the wavelet transformation allows the basic frequency of the expression to be recognized. The basic frequency indicates whether the speaker as a man or a woman.

b) Shape of the Spectral Envelope:

The spectral envelope contains information about a transfer function of the vocal tract in the articulation. The spectral envelope is dominated by the formants in a voiced region. The high-pass part of a higher transformation stage of the wavelet transformation contains this spectral envelope.

c) Spectral Tilt (Huskiness):

The huskiness in a voice is visible as negative slope in the curve of the penultimate low-pass part.

The speaker-specific characteristics a) through c) are of great significance in the speech synthesis. As initially mentioned, large sets of actually spoken expressions from which exemplary sounds are excerpted and later combined to form a new word are used in concatenative speech synthesis (synthetic speech). Discontinuities between combined sounds are thereby disadvantageous since the human ears perceives these as being unnatural. In order to oppose discontinuities, it is

advantageous to directly acquire the perceptively relevant quantities and, potentially, to compare and/or adapt them to one another.

This can occur by direct manipulation in that a speech sound is adapted at least in terms of its speaker-specific characteristics, so that it is not perceived as being
 5 disturbing in the acoustic context of the concatenatively linked sounds. It is also possible to direct the selection of a suitable sound such that speaker-specific characteristics of sounds to be linked match one another as well as possible, for example that the same or similar huskiness is inherent in the sounds.

One advantage of the invention is comprised therein that the spectral
 10 envelope reflects the articulation tract of the speaker and is not supported on formants like, for example, a pole-point model. Further, no data are lost as non-parametric representation in the wavelet transformation, the expression can always be completely reconstructed. The data proceeding from the individual transformation stages of the wavelet transformation are linearly independent of one another, can thus be influenced
 15 separately from one another and be recombined later to form the influenced expression -- loss-free.

Further, an arrangement for determining spectral speech characteristics is recited that comprises a processor unit that is configured such that an expression can be digitalized. Subsequently, the expression is subjected to a wavelet transformation
 20 and speaker-specific characteristics are determined on the basis of different transformation stages.

This arrangement is particularly suited for the implementation of the method or one of its developments explained above.

Developments of the invention also derive from the dependent claims.
 25 Exemplary embodiments of the invention are presented and explained below on the basis of the drawing.

Shown are:

Figure 1 a wavelet function;

Figure 2 a wavelet function subdivided according to real part and imaginary part;

Figure 3 a cascaded filter structure that represents the transformation steps of the wavelet transformation;

Figure 4 low-pass parts and high-pass parts of different transformation stages;

Figure 5 steps of the concatenative speech synthesis.

5 Figure 1 shows a wavelet function that is defined by

$$\psi(f) = c \cdot \left(1 - \left(\frac{f}{\sigma}\right)^2\right) \cdot e^{-\frac{1}{2} \cdot \left(\frac{f}{\sigma}\right)^2} \quad (1),$$

whereby

f references the frequency,

σ references a standard deviation, and

c references a predetermined norming constant.

10 In particular, the standard deviation σ is defined by the prescribable location of the sideband minimum 101 in Figure 1.

Figure 2 shows a wavelet function with a real part according to Equation (1) and a Hilbert transform H of the real part as imaginary part. The complex wavelet function thus derives as

$$\Psi(f) = \psi(f) + j \cdot H\{\psi(f)\} \quad (2).$$

15 The constant c from Equation (1) is employed in order to norm the complex wavelet function:

$$\int_{-\infty}^{\infty} \Psi(f) \cdot \bar{\Psi}(f) \, df = 1 \quad (3),$$

whereby $\bar{\Psi}$ references the conjugated-complex wavelet function.

Figure 3 shows the cascaded application of the wavelet transformation. A signal 301 is filtered both by a high-pass HP1 302 as well as by a low-pass TP1 305.

20 In particular, a sub-sampling thereby occurs, i.e. the plurality of values to be stored is reduced per filter. An inverse wavelet transformation assures that the original signal

301 can in turn be reconstructed from the low-pass part TP1 305 and the high-pass part HP1 304.

Filtering in the high-pass HP1 302 is separated according to real part Re1 303 and imaginary part Im 1 304.

5 Following the low-pass filter TP1 305, the signal 310 is filtered anew both by a high-pass HP2 306 as well as by a low-pass TP2 309. The high-pass HP2 306 again comprises a real part Re2 307 and an imaginary part Im2 308. Following the send transformation stage 311, the signal is filtered again, etc.

10 When a (FFT-transformed) short-time spectrum with 256 values is assumed, then eight transformation steps are implemented (sub-sampling rate: 1/2) until the signal from the last low-pass filter TP8 corresponds to the constant part.

Figure 4 shows various transformation stages of the wavelet transformation, divided according to low-pass parts (Figures 4A, 4C and 4E) and high-pass parts (Figures 4B, 4D and 4F).

15 The basic frequency of the spoken expression can be seen from the high-pass part according to Figure 4B. In addition to the fluctuations in the amplitude, a dominating periodicity in the wavelet-filtered spectrum, the basic frequency of the speaker, can be clearly recognized. On the basis of the basic frequency, it is possible to adapt predetermined expressions to one another in the speech synthesis or to define
20 suitable expressions from a data bank with predetermined expressions.

The formants of the voice signal excerpt (the length of the voice signal excerpt corresponds to about double the basic frequency) are shown as pronounced minimums and maximums in the low-pass part of Figure 4C. The formants represent resonant frequencies in the vocal tract of the speaker. The clear presentability of the
25 formants enables an adaptation and/or a selection of suitable sound components in the concatenative speech synthesis.

The huskiness of a voice can be determined in the low-pass part of the penultimate transformation stage (given 256 frequency values in the original signal:

TP7). The descent of the course of the curve between maximum M_x and minimum M_i characterizes the degree of the huskiness.

Said three speaker-specific characteristics are thus identified and can be intentionally influenced for the speech synthesis. It is thereby of particular
 5 significance that, in inverse wavelet transformation, the manipulation of a single speaker-specific characteristic influences only this; the other perceptibly relevant quantities remain unaffected. The basic frequency can thus be designationally adjusted without the huskiness of the voice being thereby influenced.

Another possible utilization is comprised in the selection of a suitable
 10 sound segment for concatenative linking with another sound segment, whereby the two sound segments were additionally recorded by different speakers in different contexts. With determination of spectral speech characteristics, a suitable sound segment to be linked can be found since, with the characteristics, criteria are known that automatically enable a comparison of sound segments to one another according to
 15 specific rules and, thus, a selection of the suitable sound segment.

Figure 5 shows steps of a concatenative speech synthesis. A data bank is produced with a predetermined set of naturally spoken language of different speakers, whereby sound segments in the naturally spoken language are identified and stored. Numerous representatives of the various sound segments of a language derive that can
 20 be accessed by the data bank. The sound segments are, in particular, phonemes of a language or a concatenation of such phonemes. The possibilities in the compilation of new words are all the greater the smaller the sound segment is. Thus, the German language comprises a predetermined set of approximately 40 phonemes that suffice for the synthesis of nearly all words of the language. Different acoustic contexts are
 25 thereby to be taken into consideration dependent on the word in which the phoneme occurs. It is then important to embed the individual phonemes into the acoustic context such that discontinuities, which human hearing senses as unnatural and "synthetic", are avoided. As mentioned, the sound segments stem from different speakers and thus exhibit different speaker-specific characteristics. In order to

synthesize an expression that has as natural an effect as possible, it is important to minimize the discontinuities. This can ensue by adaptation of the identifiable and modifiable speaker-specific characteristics or by selecting suitable sound segments from the data bank, whereby the speaker-specific characteristics likewise represent a critical aid in the selection.

By way of example, Figure 5 shows two sounds A 507 and B 508 that respectively exhibit individual sound segments 505 or, respectively, 506. The sounds A 507 and B 508 respectively derive from a spoken expression, whereby the sound A 507 is clearly distinct from the sound B 508. A parting line 509 indicates whereby the sound A 507 is to be linked to the sound B 508. In the present case, the first three sound segments of the sound A 507 are to be concatenatively linked with the last three sound segments of the sound B 508.

A temporal stretching or compression (see arrow 503) of the sound segments is implemented along the parting line 509 in order to avoid the discontinuous impression at the transition 509.

One version is comprised in an abrupt transition of the sounds parted along the parting line 509. However, said discontinuities that human hearing perceives as disturbing thereby occur. When, in contrast, a sound C is compiled [...] that the sound segments within a transition region 501 or 502 are considered, whereby a spectral distance criterion is adapted between two sound segments that can be allocated to one another in the respective transition region 501 or 502 (gradual transition between the sound segments). The Euclidean distance between the coefficients that are relevant in this region is utilized as the distance criterion, especially in the wavelet space.

Bibliography

- [1] I. Daubechies, "Ten Lectures on Wavelets", Siam Verlag, 1992, ISBN 0-89871-274-2, Chapter 5.1, pages 129-137.

Patent Claims

1. Method for determining spectral speech characteristics in a spoken expression,
 - a) whereby the expression is digitalized;
 - 5 b) whereby the digitalized expression is subjected to a wavelet transformation;
 - c) whereby the speaker-specific characteristics are defined on the basis of different transformation stages of the wavelet transformation.
- 10 2. Method according to claim 1, whereby a windowed transformation of the digitalized expression into a frequency domain is implemented before the wavelet transformation.
3. Method according to claim 2, whereby the transformation into the frequency domain is implemented with fast Fourier transformation.
- 15 4. Method according to one of the preceding claims, whereby a low-pass part and a high-pass part of a signal to be transformed are determined in each stage of the wavelet transformation.
5. Method according to one of the preceding claims, whereby a high-pass part is subdivided according to a real part and an imaginary part.
- 20 6. Method according to one of the preceding claims, whereby the wavelet transformation comprises a plurality of transformation stages, whereby the last transformation stage supplies a constant part of the expression in a repeated low-pass filtering corresponding to the plurality of transformation stages.

7. Method according to one of the preceding claims, whereby the speaker-specific characteristics are determined by

- a) a basic frequency of the spoken expression;
- b) spectral envelope;
- 5 c) a huskiness of the spoken expression.

8. Employment of the method according to one of the claims 1 through 7 for speech synthesis, whereby individual speaker-specific characteristics are adapted in view of a natural sounding concatenation of speech sounds.

9. Employment of the method according to one of the claims 1 through 7 for speech synthesis, whereby those speech sounds from a predetermined data set that assure a natural sounding concatenation of speech sounds are selected on the basis of individual spectral speech characteristics.

10. Arrangement for determining spectral speech characteristics in a spoken expression, comprising a processor unit that is configured such that the following steps can be implemented:

- a) the expression is digitalized;
- b) the digitalized expression is subjected to a wavelet transformation;
- c) the speaker-specific characteristics are defined on the basis of different transformation stages of the wavelet transformation.

Abstract**METHOD AND ARRANGEMENT FOR DETERMINING SPECTRAL SPEECH CHARACTERISTICS IN A SPOKEN EXPRESSION**

- Spectral speech characteristics are determined in a naturally spoken expression, whereby the expression is digitalized and subjected to a wavelet transformation. The speaker-specific characteristics proceed from the different transformation stages of the wavelet transformation. In the course of a speech synthesis, these characteristics can be compared to characteristics of other expressions in order to generate a synthetic speech signal that sounds continuous to the human ear.
- Alternatively, the characteristics can also be designationally modified in order to counter a perceptive dissonance.

1/3
FIG 1

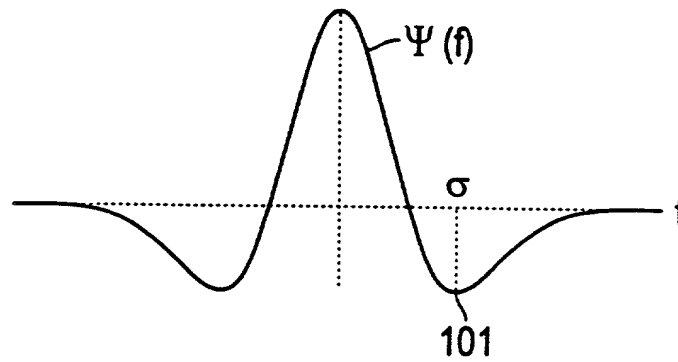


FIG 2

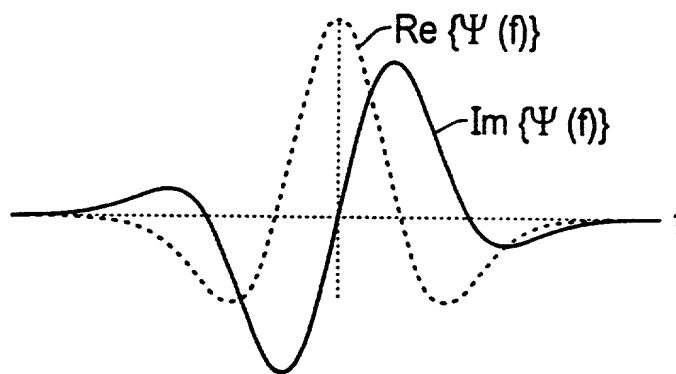
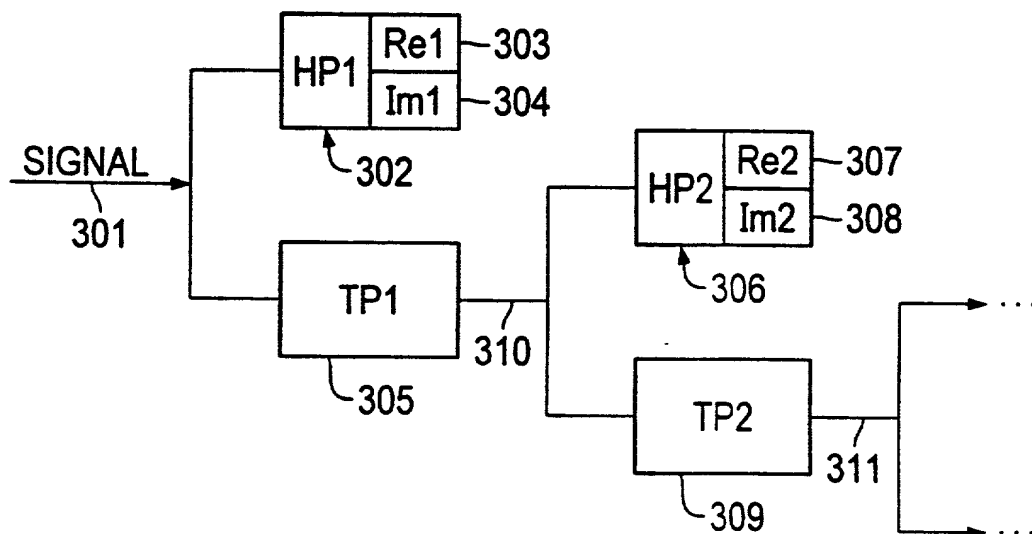
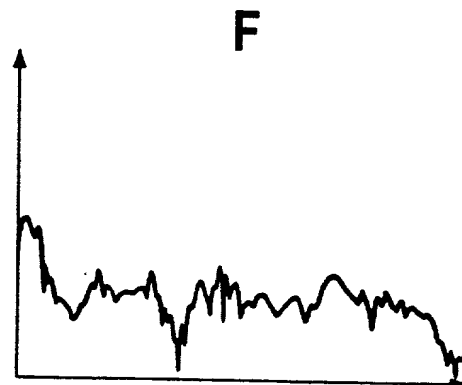
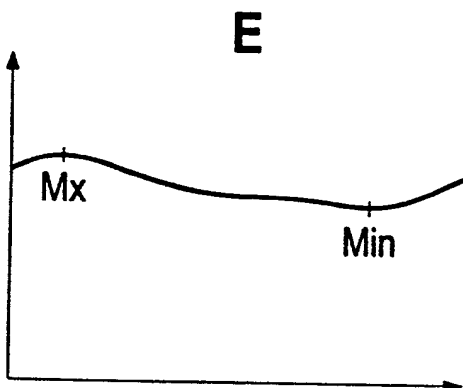
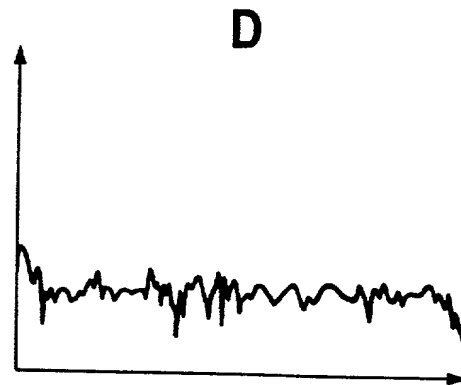
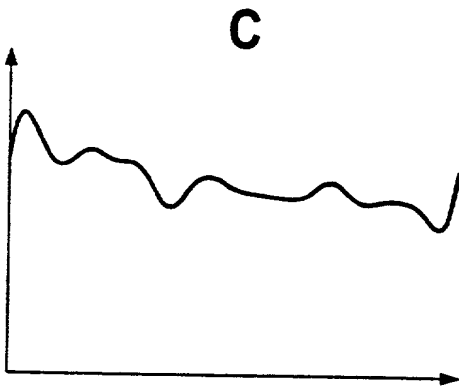
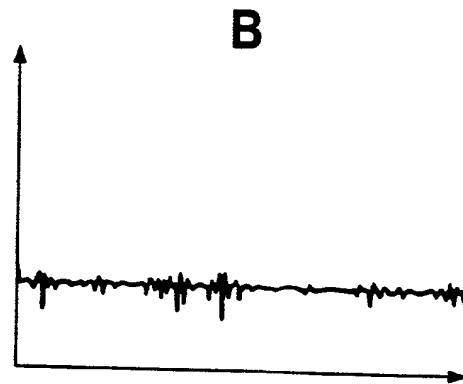
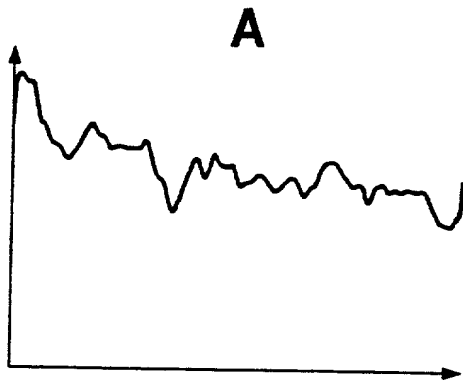


FIG 3



2/3

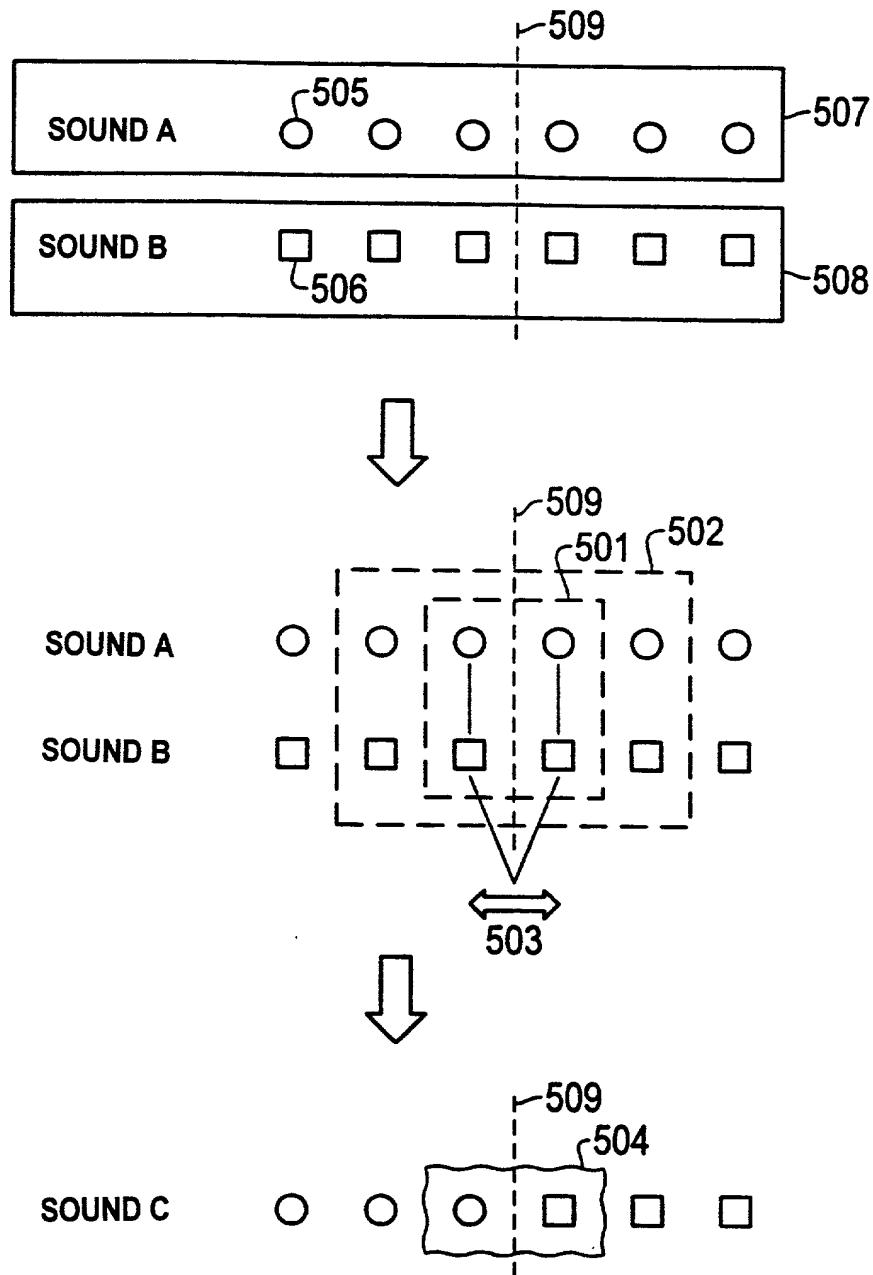
FIG 4



09/700311-4400

3/3

FIG 5



Declaration and Power of Attorney For Patent Application

Erklärung Für Patentanmeldungen Mit Vollmacht

German Language Declaration

Als nachstehend benannter Erfinder erkläre ich hiermit an Eides Statt:

dass mein Wohnsitz, meine Postanschrift, und meine Staatsangehörigkeit den im Nachstehenden nach meinem Namen aufgeführten Angaben entsprechen,

dass ich, nach bestem Wissen der ursprüngliche, erste und alleinige Erfinder (falls nachstehend nur ein Name angegeben ist) oder ein ursprünglicher, erster und Miterfinder (falls nachstehend mehrere Namen aufgeführt sind) des Gegenstandes bin, für den dieser Antrag gestellt wird und für den ein Patent beantragt wird für die Erfindung mit dem Titel:

Verfahren und Anordnung zur Bestimmung spektraler Sprachcharakteristika in einer gesprochenen Äußerung

deren Beschreibung

(zutreffendes ankreuzen)

☒ hier beigefügt ist.

☐ am _____ als

PCT internationale Anmeldung

PCT Anwendungsnummer _____

eingereicht wurde und am _____
abgeändert wurde (falls tatsächlich abgeändert).

Ich bestätige hiermit, dass ich den Inhalt der obigen Patentanmeldung einschliesslich der Ansprüche durchgesehen und verstanden habe, die eventuell durch einen Zusatzantrag wie oben erwähnt abgeändert wurde.

Ich erkenne meine Pflicht zur Offenbarung irgendwelcher Informationen, die für die Prüfung der vorliegenden Anmeldung in Einklang mit Absatz 37, Bundesgesetzbuch, Paragraph 1.56(a) von Wichtigkeit sind, an.

Ich beanspruche hiermit ausländische Prioritätsvorteile gemäss Abschnitt 35 der Zivilprozessordnung der Vereinigten Staaten, Paragraph 119 aller unten angegebenen Auslandsanmeldungen für ein Patent oder eine Erfindersurkunde, und habe auch alle Auslandsanmeldungen für ein Patent oder eine Erfindersurkunde nachstehend gekennzeichnet, die ein Anmeldedatum haben, das vor dem Anmeldedatum der Anmeldung liegt, für die Priorität beansprucht wird.

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

the specification of which

(check one)

☐ is attached hereto.

☐ was filed on _____ as

PCT international application

PCT Application No. _____

and was amended on _____
(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Variable	Mean	SD	Min	Max
Age	34.5	10.2	22	55
Gender	50%	50%	0	100
Marital status	65%	35%	0	100
Education	12.5	1.5	9	16
Income	3500	1500	1000	8000
Health status	75%	25%	0	100
Stress level	60%	20%	0	100
Life satisfaction	70%	20%	0	100
Work-life balance	65%	25%	0	100
Family support	70%	30%	0	100
Community involvement	55%	35%	0	100
Volunteering	40%	40%	0	100
Charitable giving	30%	40%	0	100
Political participation	25%	45%	0	100
Environmental activism	20%	50%	0	100
Religious participation	35%	45%	0	100
Cultural participation	45%	40%	0	100
Sports participation	50%	35%	0	100
Artistic participation	30%	50%	0	100
Music participation	40%	45%	0	100
Dance participation	35%	40%	0	100
Theater participation	25%	55%	0	100
Concert participation	30%	50%	0	100
Exhibition participation	20%	60%	0	100
Workshop participation	15%	65%	0	100
Networking	60%	40%	0	100
Professional development	55%	45%	0	100
Leadership training	45%	55%	0	100
Public speaking	35%	65%	0	100
Writing	25%	75%	0	100
Reading	70%	30%	0	100
Travel	50%	50%	0	100
Volunteering hours	10	20	0	50
Charitable donations	50	100	0	500
Political participation	20	30	0	100
Environmental activism	15	35	0	100
Religious participation	30	40	0	100
Cultural participation	40	50	0	100
Sports participation	50	60	0	100
Artistic participation	30	70	0	100
Music participation	40	60	0	100
Dance participation	35	55	0	100
Theater participation	25	75	0	100
Concert participation	30	70	0	100
Exhibition participation	20	80	0	100
Workshop participation	15	85	0	100
Networking	60	40	0	100
Professional development	55	45	0	100
Leadership training	45	55	0	100
Public speaking	35	65	0	100
Writing	25	75	0	100
Reading	70	30	0	100
Travel	50	50	0	100

Priority Claimed

<input checked="" type="checkbox"/>	<input type="checkbox"/>
Yes	No
Ja	Nein

<input type="checkbox"/>	<input type="checkbox"/>
Yes	No
Ja	Nein

☐ Yes
☐ No
Ja Nein

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §122, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application.

(Status)
(patented, pending,
abandoned)

(Status)
(patented, pending,
abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

German Language Declaration

VERTRETUNGSVOLLMACHT: Als benannter Erfinder beauftrage ich hiermit den nachstehend benannten Patentanwalt (oder die nachstehend benannten Patentanwälte) und/oder Patent-Agenten mit der Verfolgung der vorliegenden Patentanmeldung sowie mit der Abwicklung aller damit verbundenen Geschäfte vor dem Patent- und Warenzeichenamt: (Name und Registrationsnummer anführen)

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. (list name and registration number)

And I hereby appoint

Messrs. John D. Simpson (Registration No. 19,842), Lewis T. Steadman (17,074), William C. Stueber (16,453), P. Phillips Connor (19,259), Dennis A. Gross (24,410), Marvin Moody (16,549), Steven H. Noll (28,982), Brett A. Valiquet (27,841), Thomas I. Ross (29,275), Kevin W. Guynn (29,927), Edward A. Lehmann (22,312), James D. Hobart (24,149), Robert M. Barrett (30,142), James Van Santen (16,584), J. Arthur Gross (13,615), Richard J. Schwarz (13,472) and Melvin A. Robinson (31,870), David R. Metzger (32,919), John R. Garrett (27,888) all members of the firm of Hill, Steadman & Simpson, A Professional Corporation.

Telefongespräche bitte richten an:
(Name und Telefonnummer)

Direct Telephone Calls to: (name and telephone number)

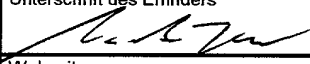
312/876-0200
Ext. _____

Postanschrift:

Send Correspondence to:

HILL, STEADMAN & SIMPSON
A Professional Corporation
85th Floor Sears Tower, Chicago, Illinois 60606

*see enclosed
change of address*

Voller Name des einzigen oder ursprünglichen Erfinders:		Full name of sole or first inventor:	
<u>HOLZAPFEL, Martin</u>			
Unterschrift des Erfinders	Datum	Inventor's signature	Date
	<u>15.8.85</u>		
Wohnsitz		Residence	
<u>D-80933 München, Germany DEX</u>			
Staatsangehörigkeit		Citizenship	
<u>Bundesrepublik Deutschland</u>			
Postanschrift		Post Office Address	
<u>Stabelerstr. 13</u>			
<u>D-80933 München</u>			
<u>Bundesrepublik Deutschland</u>			
Voller Name des zweiten Miterfinders (falls zutreffend):		Full name of second joint inventor, if any:	
Unterschrift des Erfinders	Datum	Second inventor's signature	Date
Wohnsitz		Residence	
Staatsangehörigkeit		Citizenship	
Postanschrift		Post Office Address	

(Bitte entsprechende Informationen und Unterschriften im Falle von dritten und weiteren Miterfindern angeben).

(Supply similar information and signature for third and subsequent joint inventors).

09/700311
529 Rec'd PCT/PTC 13 NOV 2000

BOX PCT
IN THE UNITED STATES DESIGNATED/ELECTED OFFICE
OF THE UNITED STATES PATENT AND TRADEMARK OFFICE
UNDER THE PATENT COOPERATION TREATY--CHAPTER II

CHANGE OF ADDRESS OF APPLICANTS' REPRESENTATIVE

APPLICANT(S): MARTIN HOLZAPFEL
ATTORNEY DOCKET NO.: P00,1796
INTERNATIONAL APPLICATION NO: PCT/DE99/01308
INTERNATIONAL FILING DATE: 03 MAY 1999
INVENTION: METHOD AND ARRANGEMENT FOR DETERMINING SPECTRAL
SPEECH CHARACTERISTICS IN A SPOKEN EXPRESSION

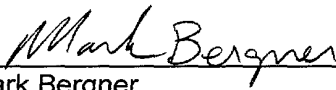
Assistant Commissioner for Patents,
Washington, D.C. 20231

S I R:

Members of the firm of Hill & Simpson designated on the original Power of Attorney have merged into the firm of Schiff Hardin & Waite. All future correspondence for the above-referenced application therefore should be sent to the following address:

SCHIFF HARDIN & WAITE
Patent Department
6600 Sears Tower
233 South Wacker Drive
Chicago, Illinois 60606-6473

Submitted by,

 (Reg. No. 45,877)
Mark Bergner
SCHIFF HARDIN & WAITE
Patent Department
6600 Sears Tower
Chicago, Illinois 60606-6473
Telephone: (312) 258-5779
Attorneys for Applicants